What is claimed is:

1. A method for modeling the availability of a cluster, the cluster having a plurality of software components and at least one node, the method comprising:

determining a plurality of component availability models using a repair model and a plurality of failure parameters, each of the plurality of component availability models corresponding to one of the plurality of software components;

combining the plurality of component availability models;

determining repair rates for node and cluster reboots; and

constructing an availability model based on the repair rates and the combined plurality of component availability models.

2. The method of claim 1, wherein the repair model includes one or more repair modes.

3. The method of claim 2, wherein the one or more repair modes of the repair model include component soft-restart, component warm-restart, component cold-restart, component fail-over, node reboot and cluster reboot.

4. The method of claim 1, wherein the plurality of failure parameters include a failure rate, repair rate and efficacy.

5. The method of claim 4, wherein the combining step further comprises:

obtaining aggregate failure rates, aggregate repair rates, and aggregate efficacies for the plurality of component availability models,

wherein the aggregate failure rates, the aggregate repair rates and the aggregate efficacies are obtained for each repair mode in the repair model.

6. The method of claim 5,

wherein for each repair mode in the repair model, an aggregate failure rate is a sum of failure rates of the plurality of software components for the repair mode,

wherein for each repair mode in the repair model, an aggregate repair rate is a weighted average of repair rates of the plurality of software components for the repair mode, weights being corresponding failure rates of the plurality of software components for the repair mode, and

wherein for each repair mode in the repair model, an aggregate efficacy is an weighted average of efficacies of the plurality of software components for the repair mode, weights being corresponding failure rates of the plurality of software components for the repair mode.

7. The method of claim 4, wherein the combining step further comprises:

for each repair mode in the repair model, aggregating failure rates of each of the plurality of software components;

for each repair mode in the repair model, aggregating repair rates of each of the plurality of software components; and

for each repair mode in the repair model, aggregating efficacies of each of the plurality of software components.

8. The method of claim 1, wherein the determining repair rates step further comprises:

specifying times that a bare platform and the cluster requires for rebooting a node and the cluster;

specifying an efficacy for node reboots;

defining cluster specific summation functions for obtaining restart times; and

combining the restart times.

9. The method of claim 1, wherein the determining the plurality of component availability models step further includes,

building an escalation graph for each of the plurality of software components.

10. The method of claim 9, wherein the escalation graph for each software component includes a weighted directed graph with its nodes representing repair modes for the software component and its edges having transition rates.

11. The method of claim 1, wherein the constructing step further comprises:

calculating a plurality of state-space parameters;

constructing a state-space model of the cluster; and

solving the state-space model.

12. The method of claim 11, wherein the plurality of state-space parameters include aggregate failure rates, aggregate repair rates, aggregate efficacies, and the repair rates for node and cluster reboots, and

wherein an aggregate failure rate, an aggregate repair rate and an aggregate efficacy is assigned to each repair mode in the repair model.

13. The method of claim 11, wherein the state-space model is represented as a weighted directed graph with its nodes representing states and its edges having transition rates.

14. The method of claim 13, wherein the states are based on the repair model.

15. The method of claim 1, wherein the plurality of component availability models include models for operation system software and models for non-operating system software.

16. A system for modeling the availability of a cluster, the cluster having a plurality of software components and at least one node, the system comprising:

means for determining a plurality of component availability models using a repair model and a plurality of failure parameters, each of the plurality of component availability models corresponding to one the plurality of software components;

means for combining the plurality of component availability models;

means for determining repair rates for node and cluster reboots; and

means for constructing an availability model based on the repair rates and the combined plurality of component availability models.

17. The system of claim 16, wherein the repair model includes one or more repair modes.

18. The system of claim 17, wherein the one or more repair modes of the repair model include component soft-restart, component warm-restart, component cold-restart, component fail-over, node reboot and cluster reboot.

19. The system of claim 16, wherein the plurality of failure parameters include a failure rate, repair rate and efficacy.

20. The system of claim 19, wherein the combining means further comprises:

means for obtaining aggregate failure rates, aggregate repair rates, and aggregate efficacies for the plurality of component availability models,

wherein the aggregate failure rates, the aggregate repair rates and the aggregate efficacies are obtained for each repair mode in the repair model.

21. The system of claim 20,

wherein for each repair mode in the repair model, an aggregate failure rate is a sum of failure rates of the plurality of software components for the repair mode,

wherein for each repair mode in the repair model, an aggregate repair rate is a weighted average of repair rates of the plurality of software components for the repair mode, weights being corresponding failure rates of the plurality of software components for the repair mode, and

wherein for each repair mode in the repair model, an aggregate efficacy is a weighted average of efficacies of the plurality of software components for the repair mode, weights being corresponding failure rates of the plurality of software components for the repair mode.

22. The system of claim 19, wherein the combining means further comprises:

for each repair mode in the repair model, means for aggregating failure rates of each of the plurality of software components;

for each repair mode in the repair model, means for aggregating repair rates of each of the plurality of software components; and

for each repair mode in the repair model, means for aggregating efficacies of each of the plurality of software components.

23. The system of claim 16, wherein the determining repair rates means further comprises:

means for specifying times that a bare platform and the cluster requires for rebooting a node and the cluster;

means for specifying an efficacy for node reboots;

means for defining cluster specific summation functions for obtaining restart times; and

means for combining the restart times.

24. The system of claim 16, wherein the determining the plurality of component availability models means further includes,

means for building an escalation graph for each of the plurality of software components.

25. The system of claim 24, wherein the escalation graph for each software component includes a weighted directed graph with its nodes representing repair modes for the software component and its edges having transition rates.

26. The system of claim 16, wherein the constructing means further comprises:

means for calculating a plurality of state-space parameters;

means for constructing a state-space model of the cluster; and

means for solving the state-space model.

27. The system of claim 26, wherein the plurality of state-space parameters include aggregate failure rates, aggregate repair rates, aggregate efficacies, and the repair rates for node and cluster reboots, and

wherein an aggregate failure rate, an aggregate repair rate and an aggregate efficacy is assigned to each repair mode in the repair model.

28. The system of claim 26, wherein the state-space model is represented as a weighted directed graph with its nodes representing states and its edges having transition rates.

29. The system of claim 28, wherein the states are based on the repair model.

30. The system of claim 16, wherein the plurality of component availability models include models for operation system software and models for non-operating system software.

31. A method for modeling the availability of a cluster, the cluster having a plurality of software components and at least one node, the method comprising:

specifying a repair model, the repair model having one or more repair modes;

specifying a plurality of failure parameters,

for each software component in the plurality of software components, assigning values to the plurality of failure parameters for each appropriate repair mode for the software component;

combining values of the plurality of failure parameters of the plurality of software components for each repair mode in the repair model;

determining repair rates for node and cluster reboots; and

constructing an availability model based on the repair rates and the combined plurality of failure parameters.

32. The method of claim 31, further comprising constructing an escalation graph for each of the plurality of software components.

33. The method of claim 31, wherein the one or more repair modes include component soft-restart, component warm-reset, component cold-restart, component fail-over, node reboot and cluster reboot.

34. The method of claim 31, wherein the plurality of failure parameters includes a failure rate, repair rate and efficacy.

35. The method of claim 31, wherein the combining step further includes:

for each repair mode in the repair model, aggregating values of each of the plurality of failure parameters.

36. A computer program product comprising a computer useable medium having computer readable code embodied therein for modeling the availability of a cluster, the cluster having a plurality of software components and at least one node, the computer program product adapted when run on a computer to effect steps including:

determining a plurality of component availability models using a repair model and a plurality of failure parameters, each of the plurality of component availability models corresponding to one of the plurality of software components;

combining the plurality of component availability models;

determining repair rates for node and cluster reboots; and

constructing an availability model based on the repair rates and the combined plurality of component availability models.